

AHG7: Analysis of HDR metrics

JVET-G0153



Currently considered HDR Luma/Luminance related

- wPSNR-Y - weighted PSNR of the Y' component in BT.2100 PQ representation, where square distortion is weighted by a luma-dependent weight similar to the scaling used for the HDR anchors using a luma-dependent QP adaptation.
- tPSNR-Y - PSNR computed from the distortion of the linear light Y component, from the CIE XYZ representation, converted to a non-linear value using the PQ reference OETF.
- PSNR-L100 - PSNR computed from the distortion of the L component from the CIE La*b* representation.

$$\text{weight}(x) = 2^{\max(-3, \min(6, 0.015 \cdot x - 1.5 - 6)) / 3}$$

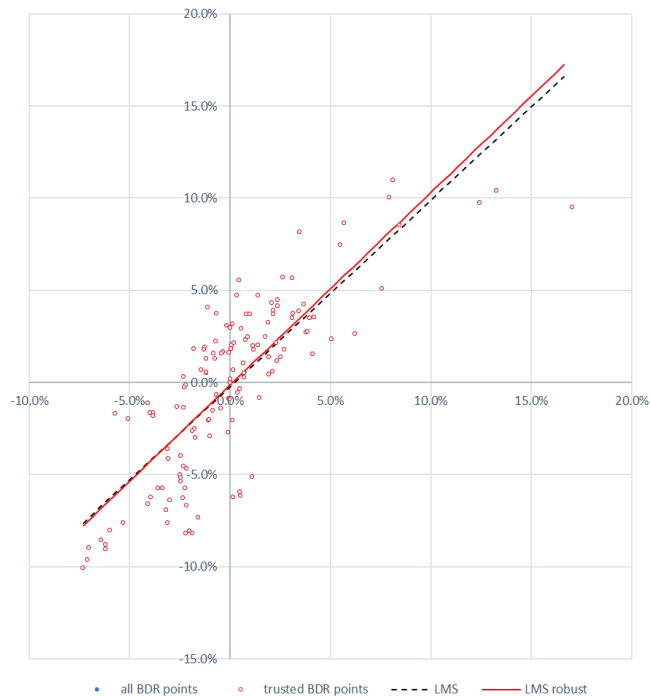
Currently considered HDR Chroma/Chrominance related

- wPSNR-U/V - similar to wPSNR-Y, but for the Cb and Cr components.
- DE100 - PSNR computed from the CIE deltaE2000 metric, derived from the signal representation in the CIE La*b* space.
- Avg_wPSNRUV - average of wPSNR-U and wPSNR-V.

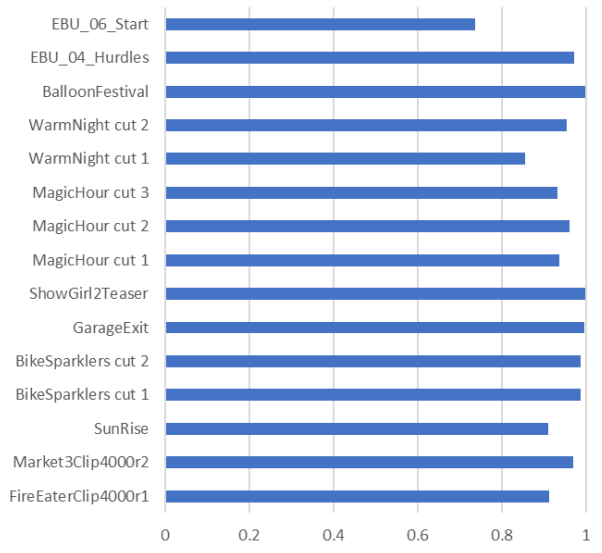
Robust linear regression parameters

correlation coef – 0.81

linear reg. $1.045 * x - 0.001$



average weight per sequence



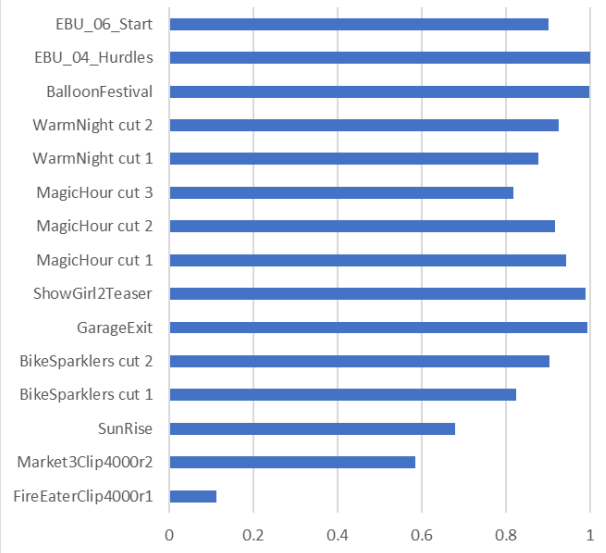
Robust linear regression parameters

correlation coef – 0.61

linear reg. $0.471 * x + 0.004$



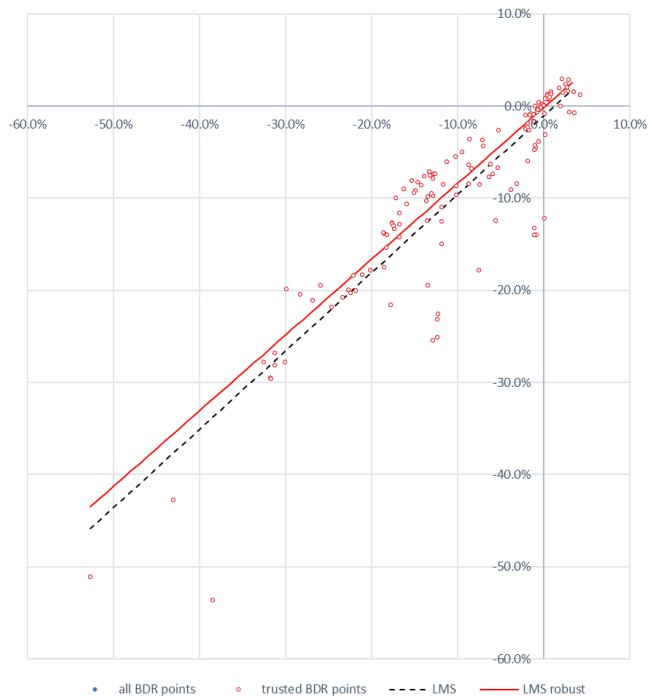
average weight per sequence



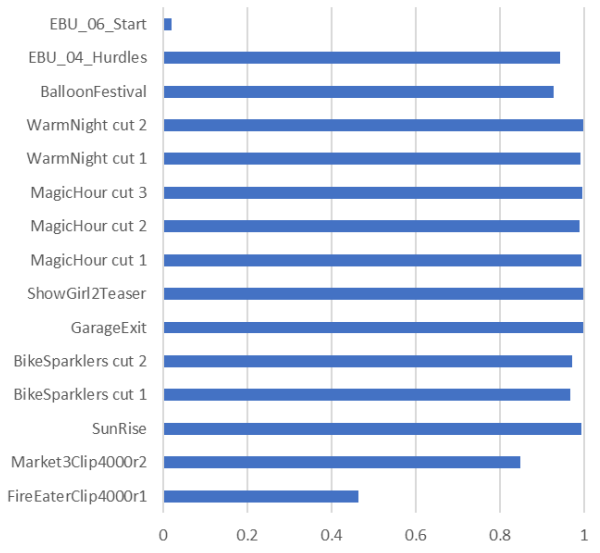
Robust linear regression parameters

correlation coef – 0.90

linear reg. $0.823 * x - 0.002$



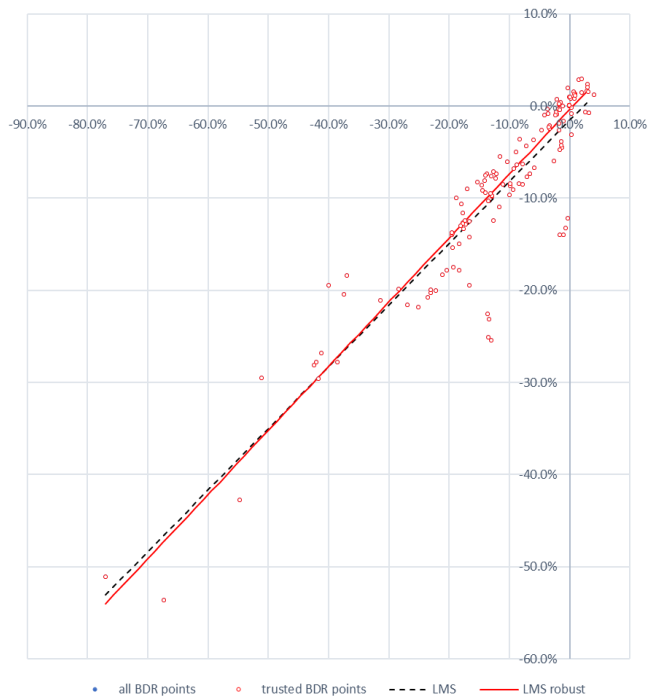
average weight per sequence



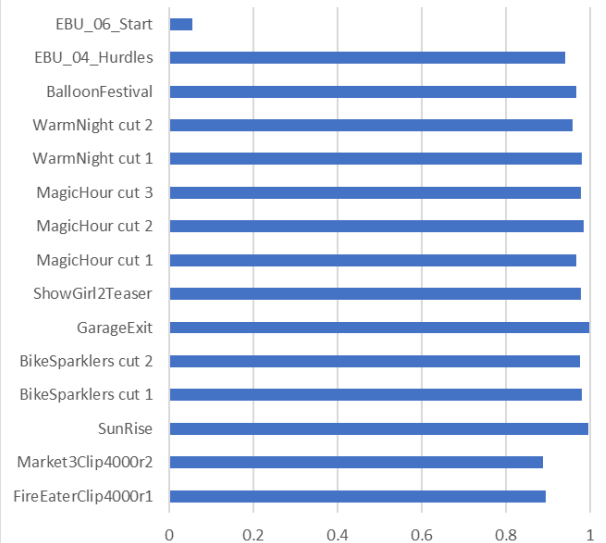
Robust linear regression parameters

correlation coef – 0.92

linear reg. $0.696 * x - 0.004$

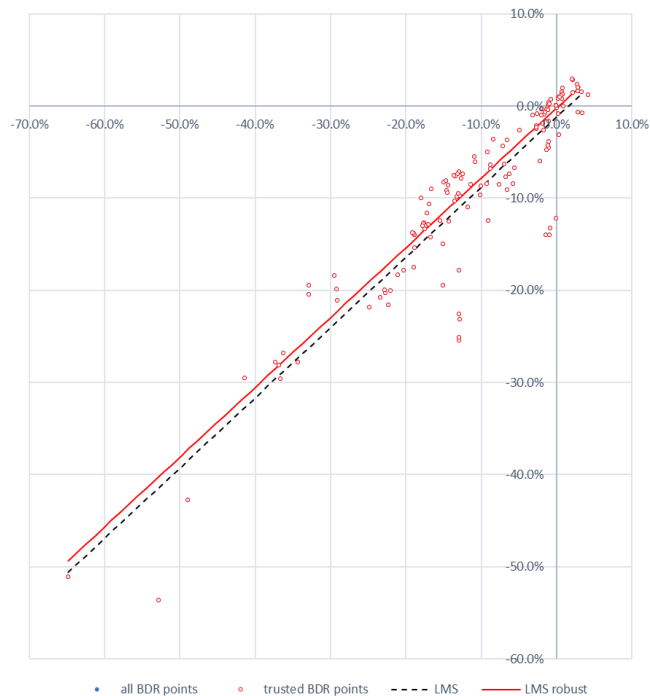


average weight per sequence

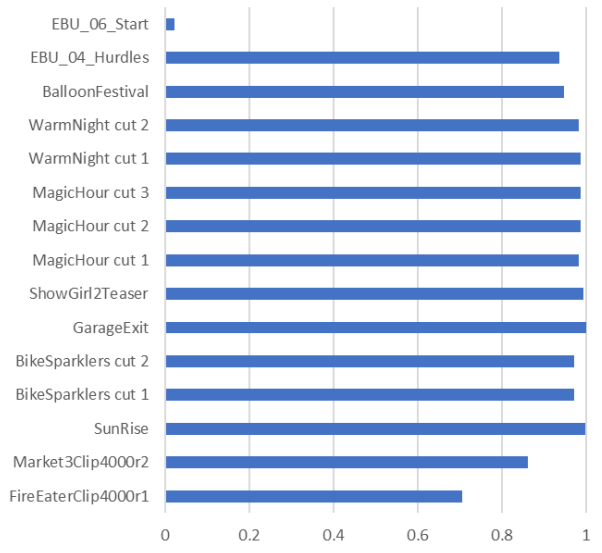


Robust linear regression parameters

correlation coef – 0.92
linear reg. $0.757 * x - 0.002$



average weight per sequence



Luma/Luminance related metrics

- wPSNR-Y vs PSNRL100 seems to be more correlated and to show better similarity trends than wPSNR-Y vs tPSNR-Y
- One possible recommendation is to reduce the Luma/Luminance related metrics to 1 primary metric (preferably wPSNR-Y or PSNRL100), but to keep the other ones as indicative metrics to detect potential problematic results when the different metrics show significantly contradictory trends.

Chroma/Chrominance related metrics

- wPSNR-U, wPSNR-V and their average, look well correlated to DE100, and show quite similar trends (except for Starting)
- It is suggested to choose one of them, either keep wPSNR-U/V or DE100